RESEARCH ON PREDICTION OF SOIL ORGANIC MATTER CONTENT BASED ON HYPERSPECTRAL IMAGING
/
土壤有机质含量的高光谱成像预测研究

Guoliang WANG[1,2]; Huiling DU[3]; Wenjun WANG[1]; Jiangui ZHAO[1]; Hong LI[2]; Erhu GUO[2]; Zhiwei LI*[1]
[1]College of Engineering, Shanxi Agriculture University, Taigu/China
[2]Millet Research Institute, Shanxi Agricultural University, Changzhi/China
[3]Department of Basic Education, Shanxi Agricultural University, Taigu/China
Tel: +86-0354-6289384; E-mail: lizhiweitong@163.com
DOI: https://doi.org/10.35633/inmateh-69-06

Keywords: *hyperspectral imaging, content of soil organic matter, numerical transformation, key bands extraction*

ABSTRACT
*Soil nutrient content is an important index to evaluate the growing environment of crops. Rapid access to soil nutrient information is an important requirement for the development of modern precision agriculture, while the detection of soil organic matter content is a necessary condition for understanding basic soil fertility and implementing crop precision cultivation. In this paper, the soil of rural fields in the southeast of Shanxi Province before sowing was taken as the research object. 111 soil samples to be tested were collected. After the process of drying, impurity removal, and grinding, the hyperspectral data of the Region of interest (ROI) of the samples were collected, and then the chemical determination of soil organic matter content was conducted. The original spectral data matrix was pretreated by numerical transformation operations, such as arithmetic mean, average deviation, 1st derivation, natural logarithm, and mixed multiplication, and a Partial least square regression (PLSR) quantitative analysis model was established. In these models, the obtained prediction set RP value under the pretreatment of F(A)*ln(AD) was the highest, reaching 0.8859. For spectral data preprocessed by F(A)*Ln(AD), the Competitive adaptive reweighted sampling (CARS) algorithm and Random frog (RF) algorithm were used to select key variables. The PLSR model was established by using F(A)*Ln(AD)&CARS data processing method and the RP value was increased to 0.9545. The prediction results can accurately reflect the real content of soil organic matter. The results of this study can provide theoretical support for the application of hyperspectral imaging technology in the determination of soil organic matter content, and provide a reference for the rapid detection of other soil components.*

摘要
土壤养分含量是评价作物生长环境的重要指标，能够快速获取土壤养分信息是发展现代精准农业重要要求，而对土壤有机质含量检测，是了解土壤基础地力，实施农作物精细化耕作必要条件。本文以晋东南地区农家田播种前土壤为研究对象，采集待测土壤样本 111 份，经晾晒、去杂及研磨等工序后采集样本兴趣区域(Region of interest，ROI)高光谱数据，随后进行土壤有机质含量化学测定。对采集到原始光谱数据矩阵进行均值、均差、求导、取自然对数及混合相乘等数值变换运算预处理，并建立土壤有机质偏最小二乘回归(Partial least square regression，PLSR)定量分析模型，其中运用均值一阶导数乘以均差值取自然对数(F(A)∗ln(AD))建立 PLSR 模型这种预处理方式，所得预测集 Rp 值最高，达到 0.8859。对 F(A)∗ln(AD)预处理下光谱数据分别采用竞争性自适应重加权采样法(Competitive adaptive reweighted sampling，CARS)及随机蛙跳(Random frog，RF)法选择特征波段处理，并建立偏 PLSR 模型，采用 F(A)∗ln(AD)&CARS 数据处理方式建立的 PLSR 模型，Rp 值提高为 0.9545，预测结果能够较为准确反映土壤有机质真实含量水平。本研究结果可以为高光谱成像技术应用于土壤有机质含量检测提供理论支撑，并为土壤其他成分快速检测提供参考。

[1] *Guoliang Wang, Ph.D. Stud.; Wenjun Wang, Assoc. Prof. Ph.D., Jiangui Zhao Ph.D. Stud.; Zhiwei Li Prof. Ph.D.;*
[2] *Guoliang Wang, Assis. Prof.; Erhu Guo, Prof.; Hong Li, Prof.;*
[3] *Huiling Du, Prof. Ph.D.*

INTRODUCTION

Soil is a complex system, consisting of gas, liquid, and solid phases, organic matter, and microorganisms. As an important carrier for crop growth, the content of soil organic matter has also become an important index to evaluate the growing environment of crops. The traditional determination of soil organic matter content is completed by chemical determination in the laboratory, which has high accuracy, but it is time-consuming, laborious, and involves high-cost *(Soil agricultural chemical analysis methods, 2000)*. With the diversified development of detection technology, to meet the basic requirements of precision agriculture for the efficient and accurate detection of soil nutrients and realize the digital research of soil nutrients information, researchers have introduced hyperspectral imaging technology into soil nutrient detection *(Shi Zhou et al., 2014)*. The spectral reflectance of the region of interest (ROI) of soil samples is accurately collected by the hyperspectral instrument, the spectral data matrix is established in the spectral dimension, and multi-channel image information is obtained in the image dimension. This information is a comprehensive reflection of the soil internal material composition and molecular structure acting on the spectrum *(Wang Wenjun et al., 2018)*. At present, hyperspectral imaging technology has become a research hotspot in agricultural research institutions at home and abroad, and it has been widely used in quality classification, component detection, and other related studies of agricultural products. *Said Nawar et al., (2016),* used different pretreatment methods combined with multiple regression models to predict soil moisture and organic matter. The results showed that the Multivariate Adaptive Regression Splines model had the best effect in predicting soil moisture and organic matter, and the coefficient of determination (R^2) could reach 0.90 and 0.85, respectively. *Xu D.Y. et al., (2018),* found that the PLSR model can be used for classification and prediction of soil composition. By comparing the prediction accuracy of principal component analysis, partial least square method, least square support vector machine and Cubist algorithm in soil total nitrogen content, *Morellos A. et al., (2016),* found that the Cubist algorithm had a better prediction result for soil total nitrogen content. *Wang Wenjun et al., (2018),* proposed that statistical parameters such as the average spectral curve, standard deviation curve, and variance curve calculated combined with multiplication and division in the preprocessing of spectral data, and a PLSR model was established for the preprocessed data. The results showed that the spectral data could effectively predict the soil total nitrogen content through numerical transformation, which provides theoretical support for the application of hyperspectral to the prediction of soil total nitrogen content. At present, the spectral data processing using hyperspectral technology to predict the content of soil organic matter is almost based on average spectral data (*Nie Zhe et al., 2019; Zhu Yuanli et al., 2021; Tang Haitao et al., 2021*). However, combined with the characteristics of hyperspectral imaging technology, there are few pretreatment methods of data conversion through spectral data.

In this study, the pre-sowing soil of "Yangfei Millet" planting experimental area in Wuxiang County, southeastern Shanxi Province was taken as the research object, and hyperspectral data from 111 soil samples were obtained. Combined with the advantages of hyperspectral imaging technology, the collected original spectral data matrix was preprocessed by numerical transformation operations, such as arithmetic mean, average deviation, 1st derivative, natural logarithm, and mixed multiplication. The characteristic variables of spectral data were selected by chemometric methods, and the PLSR quantitative analysis model of soil organic matter was established. By comparing the quantitative analysis models of soil organic matter content with different pretreatment methods, a prediction model of soil organic matter content was established, which provided theoretical support for the rapid detection of soil organic matter content.

MATERIALS AND METHODS

Collection and preparation of samples

The soil samples were collected from the "Yangfei millet" planting trial plot in Wuxiang County, southeast Shanxi Province. The field covers an area of 100 mu, which is collected before the millet is sown. About 1000 g of fresh samples are collected from the plowing layer in a random checkerboard pattern and then put into aluminum boxes. 111 samples are collected, numbered, and recorded sequentially. Samples after drying, impurity removal, grinding and sieving, and quartering sampling, are under test for hyperspectral data acquisition. On the condition of the heating, add excessive potassium dichromate into the samples to oxidize the soil organic matter, and use standard ferrous sulfate to titrate the remaining potassium dichromate. The soil organic matter could be measured with the deviation of the soil quality and the potassium dichromate consumption.

The introduction of hyperspectral instrument

In this experiment, a hyperspectral imager produced by Headwall Photonics was used to collect spectral data main working parts of the instrument include imaging equipment, a light source, a moving object platform, a computer, etc. In this test, the acquisition parameters of the near-infrared band are as follows: the wavelength range of the near-infrared hyperspectral image is 900 nm-1700 nm, the channel interval is 4.715 nm, and the step number i =172. In the near-infrared band, the relationship between step number I and wavelength λ_i is:

$$\lambda_i = 895.285 + 4.715i \quad (1)$$

where:

λ_i is the bands of the spectrum, i is number of spectral bands.

In the process of data collection, the translation speed of the object platform was adjusted to 16 mm/s, the exposure time of the camera was 0.9 ms, and the distance of the sample surface from the lens was 25 mm, so as to collect clear images. Before data collection, black and white correction is carried out on the instrument, and dark background image B and white background image W are scanned, respectively. After the operation of the instrument, the experimental image I of the tested sample is collected. After calculation according to Equation (2), the relative image X is obtained.

The above correction process was repeated for every 3 hyperspectral images.

$$X = \frac{I - B}{W - B} \quad (2)$$

In the experiment, the tested samples were put into the experimental vessels with a diameter of 3 cm and a depth of 1 cm to ensure that the tested samples were smooth and compact. Spectral data were collected for three times for each sample under test, and the data were sequentially numbered and stored.

Processing methods of the soil organic matter hyperspectral data

Within the setting coordinate range, the pixel points are selected and judged one by one, and the spectral data conforming to the setting ROI conditions are screened out to form a matrix and calculate the mean results. The spectral data were divided into training set and prediction set according to 2:1 by the Kennard-Stone algorithm (K-S). The original spectral data were pretreated by the arithmetic mean, average deviation, 1st derivation, natural logarithm, and mixed multiplication, and then established a PLSR prediction model. The Correlation coefficient (R) was used for the model evaluation. To improve the efficiency of operation and the model precision, using CARS and RF algorithms to select the key variables (*Tang Haitao et al., 2021; Chen Junyu et al., 2020*), and establish the PLSR model. Spectral data processing software mainly includes image fetching point and spectral data pretreatments developed based on Visual Basic and MATLAB 2020b (*The MathWorks, USA*).

RESULTS AND DISCUSSION

The measure of the soil organic matter

The content of soil organic matter was determined by the oxidation capacity method of potassium dichromate heated in oil bath. Each sample was divided into two times to take the average value. The statistical results of soil organic matter content were shown in Table 1.

Table 1

Statistic results of the soil organic matter

Sample Number	Min (mg/kg)	Max (mg/kg)	Mean (mg/kg)	Standard deviation (mg/kg)
111	21510	77650	31625.68	10050.5

The processing of the soil organic matter hyperspectral data

The extraction of the hyperspectral data

Since the spectral reflectance of a single point on the soil surface varies greatly, the modeling error will be great if a small number of pixel points are selected to draw spectral characteristic curves. Therefore, in order to improve the accuracy of the inversion model by combining the advantages of hyperspectral imaging technology, the sampling method shown in Fig. 1 was adopted in this study.

The selection rules for the process of extracting a large number of pixels in the ROI area were shown as (3), (4), and (5).

$$x_i = x_c - a + \Delta x, \quad i = 0,1,2... \left[\frac{2a}{\Delta x} \right] \tag{3}$$

$$y_i = y_c - b + \Delta y, \quad i = 0,1,2... \left[\frac{2b}{\Delta y} \right] \tag{4}$$

$$\frac{x_i^2}{a^2} + \frac{y_i^2}{b^2} \leq 1 \tag{5}$$

Where:

The (x_c, y_c) in (3) and (4) is the coordinate position of the ROI center, a and b are the semi-axis length of the ROI in the two-dimensional coordinate axis, Δx and Δy are the interval parameters of pixel points in the image along the coordinate axis direction respectively. The interval points collect was set to 1, [X] were round-off numbers of x and y to ensure the selected pixels were within the ROI.

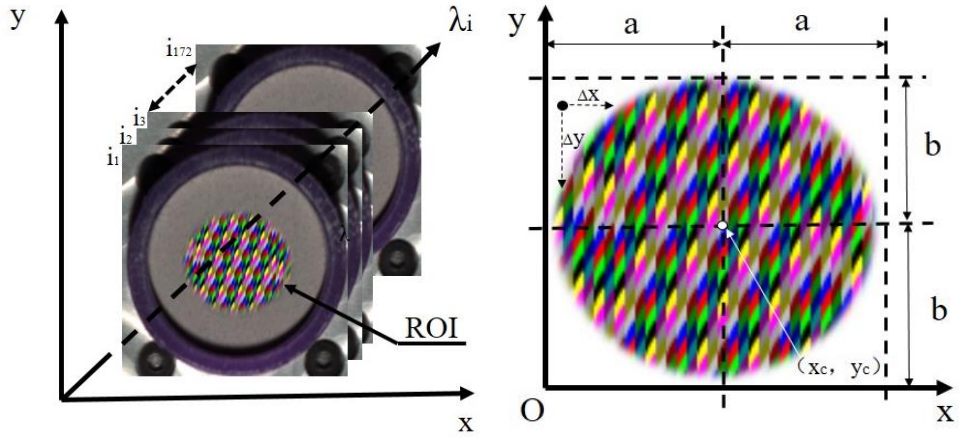


Fig. 1 - Extraction of the soil organic matter hyperspectral data

The tracking rule of pixel points was from top to bottom along the Y axis and from left to right along the X axis. The selected points were limited according to Equation (5), and 2000 fit pixel points were selected as the original spectral data.

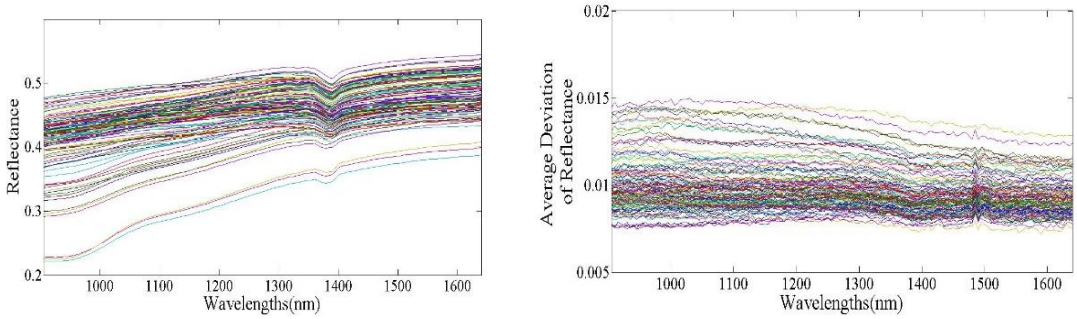
The calculation of hyperspectral data

The arithmetic mean (A) is the quotient of the sum of data divisive by the total number of data. It is widely used in the process of data statistics and analysis and has the characteristics of being sensitive, easy to be determined, and less affected by sample change (*Wang Wenjun et al., 2018*). The average deviation (AD) is the mean of the distance sum between each test value and its mean. It is sensitive and can better reflect the dispersion degree of the data to be analyzed. The calculations are shown in equations (6) and (7).

$$A_i = \frac{1}{n}(A_1 + A_2 + A_3 + \ldots + A_n) = \frac{1}{n}\sum_{i=1}^{n} A_{i,k} \tag{6}$$

$$AD_i = \frac{1}{n}\sum \left| AD_{i,k} - A_i \right| \tag{7}$$

In the equations, n is the sampling number in ROI. The number of sampling points of the near-infrared hyperspectral image is set to 2000, $A_{i,k}$ and $AD_{i,k}$ are the average spectral reflectance and the average deviation of reflectance respectively about the no. k sampling point and the no. i step length. Each step length was arranged in order to obtain the average spectral curve and average deviation spectral curve of the sample in the whole wave bands, as shown in Fig. 2.



a) b)

Fig. 2 - Spectral curves of the soil organic matter

a: Average spectral curves of the soil organic matter;
b: Average deviation spectral curves of the soil organic matter

At both ends of the near spectral range (<900 nm and >1700 nm), the spectral curve has a large disturbance, so the intercepted spectral range is 950-1650 nm, and the number of bands is reduced to 148. Soil organic matter has a variety of functional groups of organic compounds. Its spectral characteristics mainly describe the vibration in the molecules and lattice and the electron energy level transition in the atom of the high-molecular compound, spectral curve appears obvious wave was mainly caused by double frequency and co-frequency of the group like O-H and C-H in the organic matter. Figure 2 (a) shows that 1400 nm was the water absorption zone, taken as the boundary. From 950 nm-1400 nm, the spectral reflectance curve of the sample showed a monotonous upward trend and the curve shape was close to a straight line. From 1400 nm-1650 nm, the curve changed gently with a smaller slope than before (*Said Nawar et al., 2016*).

PLSR model for soil organic matter based on different pretreatments

The 1st derivative operation of the average spectral data can remove the linear part of the data and highlight the extreme value of the spectral data, which is conductive to observing the peak and valley characteristics of the data. The LN function has a good gain within the defined interval (0, 1]. The LN function operation of average spectral data can highlight the variation characteristics of spectral curves, enhance the difference of extreme values among data, and improve the accuracy of spectral feature recognition (*Wang Hongbo et al., 2017*). Table 2 lists 15 preprocessing methods results of spectral data A and AD that are calculated in 1st Derivative and LN functions, establishes PLSR model, obtains RP value of PLSR model under different pretreatments, and determines the best regression model. Among these methods, the highest RP value was based on the pretreatment F(A)*ln(AD), reaching 0.8859. The fit of the training set and prediction set was shown in Fig.3.

Table 2

PLSR modeling results of methods based on different pretreatments

Pretreatment	LVs	Calibration		Prediction	
		RMSEC (mg/kg)	R_C	RMSEP (mg/kg)	R_P
A	3	3388.29	0.9159	2793.53	0.6495
A*AD	3	6006.64	0.8271	4316.53	0.8691
A*F(A)	4	4172.7	0.9315	3897.68	0.7668
A*ln(A)	3	3518.65	0.9514	3455.46	0.7870
A*F(AD)	5	777.738	0.9898	14699.8	0.3300
A*ln(AD)	3	3518.65	0.9514	3455.46	0.7869
F(A)*AD	4	3278.6	0.9589	2864.37	0.8355
F(A)*F(AD)	1	16475.1	0.2085	6675.72	0.2396
F(A)*ln(AD)	**3**	**3417.68**	**0.9548**	**3037.72**	**0.8859**
ln(A)*AD	4	4344.2	0.9284	3553.94	0.1880
ln(A)*F(AD)	2	5147.83	0.8020	4478.75	0.4411
ln(A)*ln(AD)	3	3162.88	0.9612	3272.82	0.6675

Table 2
(continuation)

Pretreatment	LVs	Calibration		Prediction	
		RMSEC (mg/kg)	R_C	RMSEP (mg/kg)	R_P
AD	7	1669.34	0.9892	6992.65	0.4736
AD*F(AD)	1	9399.79	0.5894	2823.27	0.2283
AD*ln(AD)	3	6662.1	0.8073	4094.64	0.8136

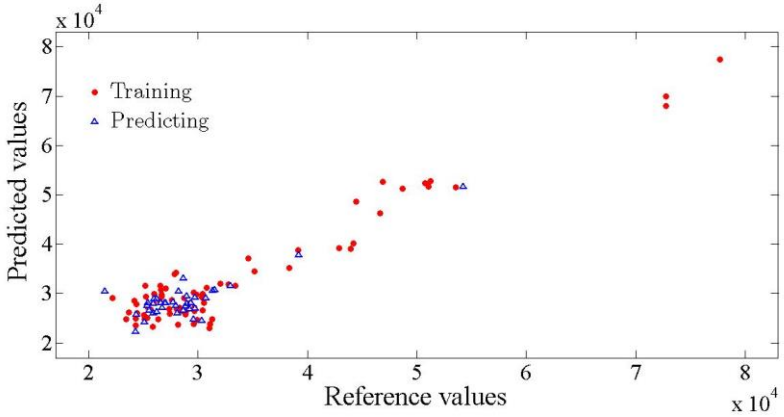


Fig. 3 - The fit of training set and prediction set by F(A)*ln(AD) pretreatment

Selection of key variables

Hyperspectral data has a large number of bands and high information redundancy, so it is easy to produce dimensionality disaster phenomena in the data processing, which leads to the decline of prediction accuracy. In this study, two key bands extraction methods, CARS and RF, were used to extract the feature band from the hyperspectral data of soil organic matter. The dimension of the hyperspectral data was reduced while the relatively complete spectral information was retained.

Selection of key variables based on CARS algorithm

Using Darwin's evolution theory as a reference, CARS algorithm selects the wavelength points with large absolute value of regression coefficient in PLSR model through adaptive reweighted sampling (ARS) technology, removes the wavelength points with small weight, selects the subset with the lowest root mean square error by cross-validation, and finds out the optimal variable combination. The algorithm was used to process spectral data and quickly select key spectral variables of samples. Fig. 4 shows the process of selecting key variables by CARS algorithm and the change curve of relevant parameters (*Tang Haitao et al., 2021*).

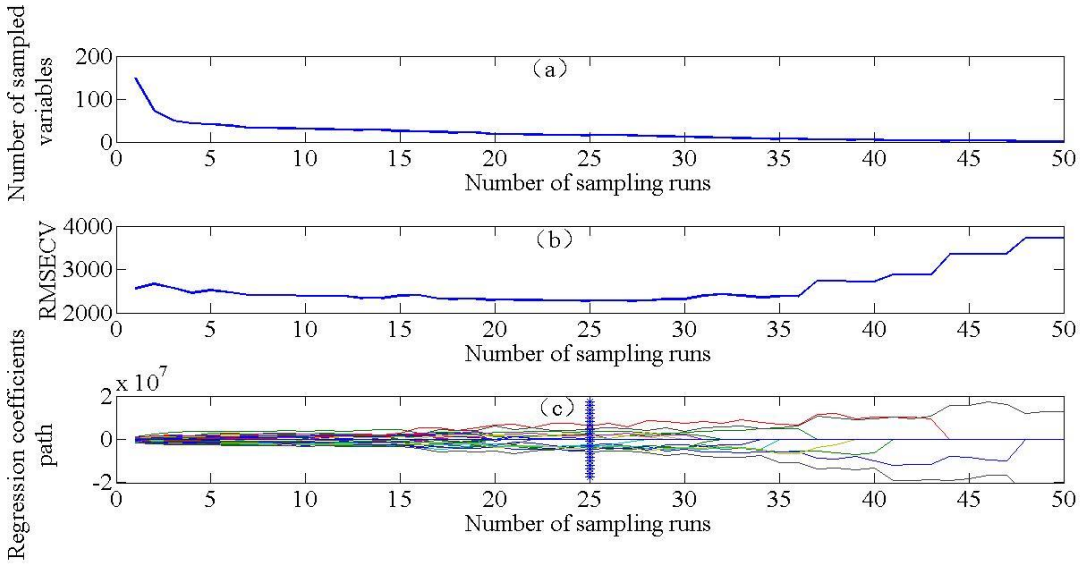


Fig. 4 - Selection of key variables using CARS algorithm
a: Changing trend of the number of sampled variables;
b: The variation of root-mean-square error of cross-validation values;
c: Regression coefficients of each variable with the increasing of sampling runs

Fig. 4(a) shows the trend of the number of samples tested with the increase of sampling times. Fig. 4(b) shows the change of Root mean square error (RMSECV) with the increase of sampling times through Cross validation (CV), with a trend of smooth decline from 1-25, 26-50 gradually rises and the minimum appears at 25, at which point RMSECV appears at its minimum. As can be seen from Fig. 4(c), the optimal selection of regression coefficients (RC) appears at the position marked with blue asterisk vertical line in the figure, where variable information can be retained as much as possible and the key variables can be selected. After the selection of key variables using CARS algorithm, as shown in Table 3, a total of 15 key variables were selected. The triple frequency of O-H appeared near 980 nm, the double frequency and co-frequency of C-H appeared near 1220 nm, and water molecular absorption peak appeared near 1400 nm.

Selection of key variables based on RF algorithm

With statistical ideas as the theoretical background, RF algorithm calculates the selection probability of each variable through multiple iterative operations. This algorithm has the advantage of not being easy to overfitting. Parameters were set before the RF algorithm. According to relevant references (*Chen Jianyu et al., 2020*), 0.4 was selected as the threshold value of the key variable. The execution was conducted for 50 times, and the running results were averaged. As shown in Fig. 5, some values in selection probability (SP) of some bands were high, and these wave peaks were highly correlated with the spectral characteristics of soil organic matter. Finally, 10 key variables were selected.

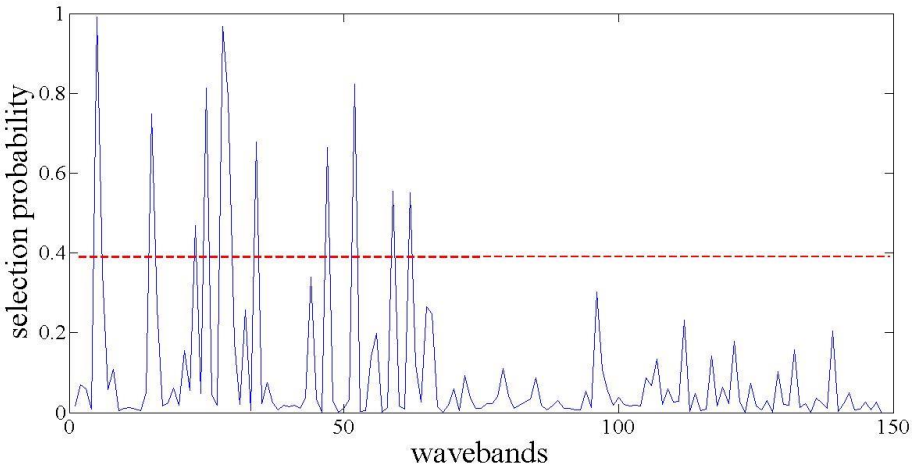


Fig. 5 - Selection probabilities of each wavelength using RF algorithm

As shown in Table 3, the key variables selected by RF and CARS algorithm almost overlap. In order to reduce the influence of random factors on the selection of key variables and ensure the effectiveness of the calculation process, the execution times of both algorithms were set at 50 times.

Table 3

Results of CARS&RF based on key wavelengths extraction

Pretreatment	Number of variables	Variables
CARS	15	959, 974, 1021, 1025, 1058, 1068, 1082, 1086, 1110, 1172, 1195, 1228, 1242, 1403, 1412
RF	10	974, 1021, 1068, 1082, 1086, 1110, 1172, 1195, 1228, 1242

Table 4 lists PLSR models of soil organic matter established by two data pretreatments, F(A)*Ln(AD)&CARS and F(A)*Ln(AD)&RF. As can be seen from the results, the RP value of PLSR established by using F(A)*Ln(AD)&CARS data pretreatments was 0.9545, while the RP value of PLSR established by using F(A)*Ln(AD)&RF decreased to 0.8834. This result may be due to the loss of relevant data in the process of using RF algorithm to extract key variables. It can be concluded from the above results that the PLSR model established by F(A)*Ln(AD)&CARS can be used to predict the hyperspectral reflectance of soil organic matter.

Table 4

PLSR modeling results of different methods based on key wavebands extraction

Pretreatment	LVs	Calibration		Prediction	
		RMSEC (mg/kg)	R_C	RMSEP (mg/kg)	R_P
F(A)*ln(AD)&CARS	3	4084.77	0.9208	3368.47	0.9545
F(A)*ln(AD)&RF	3	5128.9	0.9063	3880.41	0.8834

CONCLUSIONS

The hyperspectral data of 111 soil samples were obtained by taking the pre-sowing soil in "Yangfei millet" planting experimental area in Wuxiang County, southeast Shanxi Province as the research object.

Combined with the advantages of hyperspectral imaging technology, the collected original spectral data matrix was pretreated by numerical transformation operations, such as arithmetic mean, average deviation, 1st derivation, natural logarithm, and mixed multiplication, and the characteristic variables of spectral data were selected by the chemometric method, and then PLSR models were established respectively to realize the detection of soil organic matter. The main conclusions were as follows:

(1) Soil samples showed peaks and valleys in a specific spectral absorption zone, in which the water absorption zone was around 1400 nm and was taken as the boundary. Between 950 nm-1400nm, the spectral reflectance curve of the samples showed a monotonous upward trend and the curve shape was close to a straight line. From 1400nm-1650nm, the curve changes gently and the slope was smaller than before.

(2) The original spectral data were calculated by arithmetic mean, average deviation, 1st derivation, and natural logarithm, and then established the PLSR model of soil organic matter. In these PLSR models, the obtained prediction set R_P value under the pretreatment of F(A)*ln(AD) was the highest, reaching 0.8859.

(3) To further improve the precision of soil spectral inversion and reduce the amount of data computation, CARS and RF algorithms were used to extract the key variables from the spectral data preprocessed by F(A)*Ln(AD) and establish the PLSR model. The results showed that the PLSR model established by F(A)*Ln(AD)&CARS data pretreatment had a higher R_P value of 0.9545.

This study shows that the method of using the original spectral data arithmetic mean, average deviation, 1st derivation, and natural logarithm combined with the chemometric methods is feasible for the application of hyperspectral imaging technology in the detection of soil organic matter content and can provide a theoretical reference for the rapid detection of other soil components.

ACKNOWLEDGEMENT

This work was supported by Key Research and Development Program of Shanxi Province (No. 201903D211005), China Agriculture Research System of MOF and MARA, Technology Innovation Project of Shanxi Agricultural University (No. 2017YJ12), Shanxi Province Excellent Doctoral Work Award Fund (No. SXYBKY2019018).

REFERENCES

[1] Blasch Gerald, Spengler Daniel, Hohmann Christian, et al., (2015), Multitemporal soil pattern analysis with multispectral remote sensing data at the field-scale. *Computers and Electronics in Agriculture*, Vol. 2015, Issue 113, pp. 1-13, Oxford/England.

[2] Chen Jianyu, Li Guanghui, (2020), Prediction of moisture content of wood using Modified Random Frog and Vis-NIR hyperspectral imaging. *Infrared Physics & Technology*, Vol. 2020, Issue 105, Amsterdam/Netherlands, https://doi.org/10.1016/j.infrared.2020.103225.

[3] Gu Xiaohe, Wang Yancang, Sun Qian, et al., (2019), Hyperspectral inversion of soil organic matter content in cultivated land based on wavelet transform. *Computers and Electronics in Agriculture*, Vol. 2019, Issue 167, Oxford/England, https://doi.org/10.1016/j.compag.2019.105053.

[4] He Shaofang, Shen Luming, Xie Hongxia, (2021), Hyperspectral Estimation Model of Soil Organic Matter Content Using Generative Adversarial Networks. *Spectroscopy and Spectral Analysis*, Vol. 41, Issue 6, pp. 1905-1911, Beijing/China.

[5] Jin Xiuliang, Song Kaishan, Du Jia, et al., (2017), Comparison of different satellite bands and vegetation indices for estimation of soil organic matter based on simulated spectral configuration. *Agricultural and Forest Meteorology*, Vol. 2017, Issue 244, pp. 57-71, Amsterdam/Netherlands.

[6] Morellos A., Pantazi X.E., Moshou D., et al., (2016), Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. *Biosystems Engineering*, Vol. 2016, Issue 137, pp. 340–349, San Diego/U.S.A.

[7] Nie Zhe, Li Xiufeng, Lv Jiaxin, et al., (2019), Hyperspectral retrieval of surface soil organic matter content in a typical black soil region of northeast China. *Chinese Journal of Soil Science*, Vol. 50, Issue 6, pp. 1285-1293, Shenyang/China.

[8] Nowkandeh S.N., Noroozi A.A., Homaee., (2018), Estimating soil organic matter content from Hyperion reflectance images using PLSR, PCR, MinR and SWR models in semi-arid regions of Iran. *Environmental Development*, Vol. 2018, Issue 25, pp. 23-32, Amsterdam/Netherlands.

[9] Said Nawar, Henning Buddenbaum, Joachim Hill, et al., (2016), Estimating the soil clay content and organic matter by means of different calibration methods of vis-NIR diffuse reflectance spectroscopy. *Soil & Tillage Research*, Vol. 2016, Issue 155, pp. 510–522, Amsterdam/Netherlands.

[10] Samuel O Ihuoma, Chandra A Madramootoo, (2020), Narrow-band reflectance indices for mapping the combined effects of water and nitrogen stress in field grown tomato crops. *Biosystems Engineering*, Vol. 2020, Issue 192, pp. 133-143, Sandiego/U.S.A.

[11] Shi Zhou, Wang Qianlong, Peng Jie, et al., (2014), Development of a national VNIR soil-spectral library for soil classification and prediction of organic matter concentrations. *Science China: Earth Sciences*, Vol. 44, Issue 5, pp. 978-988, Beijing/China.

[12] Tang Haitao, Meng Xiangtian, Suo Xunxin, et al., (2021), Hyperspectral prediction on soil organic matter of different types using CARS algorithm. *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 37, Issue 2, pp. 105-113, Beijing/China.

[13] Wang Wenjun, Wang Can, Li Zhiwei, et al., (2018), Prediction of total nitrogen content in brown soil based on hyperspectral technology. *Journal of Shanxi Agricultural University (Nature Science Edition)*, Vol. 38, Issue 9, pp. 71–76, Shanxi/China.

[14] Wang Hongbo, Zhao Ziqi, Lin Yi, et al., (2017), Leaf Area Index Estimation of Spring Maize with Canopy Hyperspectral Data Based on Linear Regression Algorithm. *Spectroscopy and Spectral Analysis*, Vol. 37, Issue 5, pp. 1489-1496, Beijing/China.

[15] Xu D.Y., Ma W.Z., Chen S.C., et al., (2018), Assessment of important soil properties related to Chinese Soil Taxonomy based on vis-NIR reflectance spectroscopy. *Computers and Electronics in Agriculture,* Vol. 2018, Issue 144, pp. 118–129, Oxford/England.

[16] Zhu Yuanli, Wang Dongyan, Zhang He, et al., (2021), Soil organic carbon content retrieved by UAV-borne high resolution spectrometer. *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 37, Issue 6, pp. 66-72, Beijing/China.

[17] ***Soil science society of China, (2000), *Soil science society of China*. China Agricultural Science and Technology Press, Beijing/China.